On Hardness of Pricing Items for Single-Minded Bidders*

Rohit Khandekar Tracy Kimbrel Konstantin Makarychev Maxim Sviridenko

IBM T.J.Watson Research Center

Abstract

We consider the following *item pricing* problem which has received much attention recently. A seller has an infinite numbers of copies of n items. There are m buyers, each with a budget and an intention to buy a fixed subset of items. Given prices on the items, each buyer buys his subset of items, at the given prices, provided the total price of the subset is at most his budget. The objective of the seller is to determine the prices such that her total profit is maximized.

In this paper, we focus on the case where the buyers are interested in subsets of size at most two. This special case is known to be APX-hard (Guruswami et al [7]). The best known approximation algorithm, by Balcan and Blum, gives a 4-approximation [2]. We show that there is indeed a gap of 4 for the combinatorial upper bound used in their analysis. We further show that a natural linear programming relaxation of this problem has an integrality gap of 4, even in this special case. Then we prove that the problem is NP-hard to approximate within a factor of 2 assuming the Unique Games Conjecture; and it is unconditionally NP-hard to approximate within a factor 17/16. Finally, we extend the APX-hardness of the problem to the special case in which the graph formed by items as vertices and buyers as edges is *bipartite*.

We hope that our techniques will be helpful for obtaining stronger hardness of approximation bounds for this problem.

*{rohitk, kimbrel, konstantin, sviri}@us.ibm.com.

1 Introduction

Many pricing questions in the IT industry stem from a specific cost structure: high fixed cost of production, but near-zero or zero variable cost of production. This cost structure characterizes a class of technology products which are collectively termed *digital goods*. Put differently, the cost of producing the first unit of a digital good is very high, but the cost of producing each additional unit is virtually zero. For instance, Microsoft spends hundreds of millions of dollars on developing each version of its Windows operating system. Once this first copy of the OS has been developed, however, it can be replicated at no cost. Other examples of digital goods are pay-per-view television programs, downloadable audio files, etc.

In this paper, we consider a problem of pricing digital goods that has received a lot of attention in the computer science community recently. Consider a *monopolistic* market with a single seller who has n digital goods to sell. Since the variable cost of production is near-zero, we assume that the seller has infinite copies of each good. Suppose that there are m buyers, each buyer i associated with a fixed budget $b_i > 0$, which is the maximum amount of money he is willing to spend. Each buyer is interested in buying some bundles of digital goods. For example, a buyer may be interested in buying an operating system together with an anti-virus software; but he may not be interested in buying them separately.

We further focus on the case where each buyer is interested in exactly one subset of goods. This setting is often referred to as a market with *single-minded* buyers. While this assumption may seem unnatural, it turns out that even this special case is computationally hard for the optimization problem we consider. The seller, who is assumed to know the demand and budget information, is then posed with the following problem of pricing goods. The seller must set a price $p_j \geq 0$ for each good j — she is not allowed to price the same item differently for different buyers. For a subset S of goods, let $p(S) = \sum_{j \in S} p_j$ denote the total price of goods in S. Once the prices are fixed, each buyer i buys his subset S_i of items if its total price is at most his budget, i.e., $p(S_i) \leq b_i$. If a buyer i satisfies this condition, he pays $p(S_i)$ to the seller. If on the other hand, this condition is not satisfied, buyer i buys nothing and pays nothing to the seller. In such a model, a natural objective for the seller is to price the items so as to maximize the total profit generated, i.e., to find prices $\{p_j\}$ so as to maximize

$$\sum_{i: p(S_i) \leq b_i} p(S_i).$$

1.1 Related work

The problem of profit-maximizing pricing of goods in unlimited supply was introduced by Goldberg, Hartline, Karlin, Saks, and Wright [6]. In their setting, the buyers were interested in single goods and hence the optimization problem was trivial, and they focused on designing truthful mechanisms to maximize profit. There has been a lot of subsequent work on this and related models — below, we briefly survey only those results that are directly relevant to the problem we consider.

Guruswami, Hartline, Karlin, Kempe, Kenyon, and McSherry [7] considered the problem of profit maximization in a variety of settings, including single-minded bidders. They showed a logarithmic approximation guarantee and APX-hardness for the profit maximization problem. For single-minded bidders, a polylogarithmic hardness result was obtained by Demaine, Feige, Hajiaghayi, and Salavatipour [4]. The problem of the single-minded bidder case, where the size of the bundles demanded by the buyers was at most k, was considered by Briest and Krysta [3] who gave an $O(k^2)$ approximation for the problem, and was improved by Balcan and Blum [2] to $O(k)$. For the special case of $k = 2$, they obtain a 4-approximation algorithm.

The case of $k = 2$ (also called as the *graph pricing* problem) can be thought of as the following graph problem with goods as vertices and buyers as edges. Consider an undirected graph on n vertices and m edges. There may be parallel edges and loops. Each edge e has a budget $b_e \geq 0$. Given prices $p_v \geq 0$ on the vertices v, an edge $e = (u, v)$ is *satisfied* if $p_u + p_v \leq b_e$. The goal is to set the prices to maximize the total profit generated: $\sum_{e=(u,v)\in E: p_u + v \leq b_e}(p_u + p_v)$. The 4-approximation algorithm of Balcan and Blum [2] for this case first reduces the problem to the case where G is a bipartite graph by losing a factor of 2 in the approximation. It then gives a 2-approximation on the bipartite graphs. Recently, Krauthgamer, Mehta, and Rudra [11] focused on the case $k = 2$ with further restriction that the budgets b_e are same for all the edges; but the graph may have self-loops. In such a case, they gave an LP-rounding algorithm that yields an approximation of $\frac{6+\sqrt{2}}{5+\sqrt{2}} \approx 1.15$. They also showed a matching integrality gap for these instances.

If we assume that the goods that are being sold are the edges of a graph and that buyers are purchasing paths in this graph, we can interpret this as the problem of pricing network connections, street segments (therefore termed the tollbooth problem [7]), or other types of transportation links (e.g., railway or flight connections). If the underlying graph is just a line, then this problem is called the highway problem [7]. Interestingly, even this very restricted variant turns out to be intriguingly complex [3, 5, 7]. Hartline and Koltun [9] have presented a near-linear-time FPTAS for the practically relevant case that the number of goods for sale is a fixed constant.

1.2 Our results and techniques

In this paper, we focus on the graph pricing problem described above. The bundles of the buyers have at most two goods each, i.e., $k = 2$.

We first prove that the problem is hard to approximate within a factor of 2 assuming the Unique Games Conjecture, and within a factor of 17/16 assuming $P \neq NP$. To this end, we introduce a new problem which we call the Restricted Maximum Acyclic Subgraph problem: we are given a directed graph and our goal is to arrange its vertices on the real line so as maximize the number of forward edges. However, unlike the Maximum Acyclic Subgraph problem, we can place every vertex v only in a specified set of positions S_v (see Section 2 for details). We show that the Graph Pricing problem is at least as hard to approximate as the Restricted Maximum Acyclic Subgraph problem (in Section 3). This immediately gives us a lower bound of 2, since Restricted Maximum Acyclic Subgraph is a special case of Maximum Acyclic Subgraph, which as was recently shown by Guruswami, Manokaran, Raghavendra [8], is hard to approximate within a factor of 2 assuming the Unique Games Conjecture.

The Restricted Maximum Acyclic Subgraph problem is also a special case of MAX DICUT on directed acyclic subgraphs. We can show that MAX DICUT on directed acyclic subgraphs is at least as hard to approximate as MAX CUT. (We omit the proof from this extended abstract.) This gives us an unconditional NP-hardness of 17/16. The inapproximability of MAX CUT was established by Håstad [10].

Then we initiate a study of several algorithmic approaches that might improve the approximation guarantee. Note the following trivial upper bound on the value of the optimal solution: allow each node v to collect its maximum profit $R(v)$ from the incident edges assuming that all its neighbors are priced at 0. The overall upper bound on the optimum solution is then $\sum_v R(v)$. This observation was used by Balcan and Blum [2] in their approximation algorithm that computes a solution of value at least $\frac{1}{4}\sum_v R(v)$, and thus gives a 4-approximation. It was not known however whether this analysis of the algorithm could be improved. We show that this upper bound indeed has a gap of 4. Therefore new upper bounds are required to get a better approximation factor.

A natural linear programming relaxation (LP) gives such an upper bound. This linear program can be thought of as a generalization of the one used by Krauthgamer, Mehta and Rudra [11] to the case of arbitrary budgets. Unfortunately, it turns out that this LP also has an integrality gap of 4. The proof again uses our reductions from Restricted Maximum Acyclic Subgraph and MAX DICUT on directed acyclic subgraphs. We take a directed acyclic graph $G = (V, A)$ in which every directed cut contains at most a $(1/4 + o(1))$ fraction of all edges. (A family of such graphs was recently constructed by Alon, Bollobàs, Gyàrfàs, Lehel, and Scott [1].) We show how to transform G to an instance of the Graph Pricing problem whose solutions correspond to directed cuts in G. Therefore, every combinatorial solution to this instance has value at most $(1/4 + o(1))|A|$. Meanwhile, there is an LP solution that collects a profit 1 from every edge, and thus has value $|A|$. We describe this transformation and its analysis in Section 3.2.

Finally, we analyze the bipartite case. Note that if we improved the algorithm for bipartite graphs, we would get an improvement over the 4-approximation of Balcan and Blum for general graphs. In particular, if we could solve the problem for bipartite graphs exactly, we would get a 2-approximation for general graphs. Unlike the general case of the graph pricing problem, the bipartite case was not even known to be NP-hard. We show that it is in fact APX-hard by a reduction from MAX CUT. We present the proof in the Appendix.

2 Preliminaries

Let us fix some notation. An instance of the Graph Pricing problem $\Pi = (G, b)$ is a pair consisting of a graph $G = (V, E)$ and a set of budgets $\{b_e\}$, $e \in E$. Throughout the paper we assume that the budgets are positive integers and that the graph does not have parallel edges or self-loops. A solution of the problem is an arbitrary assignment of prices to the vertices, i.e., a set of nonnegative real numbers $\{p_v\}_{v\in V}$. The profit of the solution is

$$\text{profit}_\Pi(p) = \sum_{e=(u,v)\in E} \begin{cases} p_u + p_v, & \text{if } p_u + p_v \le b_e; \\ 0, & \text{otherwise.} \end{cases}$$

We denote the profit of the optimal solution by OPT_Π:

$$OPT_\Pi = \max_{p_v \in \mathbb{R}^+ \cup \{0\}} \text{profit}_\Pi(p).$$

In the proof we consider a more general version of the Graph Pricing problem, in which the graph may have parallel edges and edges are weighted. We denote the weight of an edge e by w_e. We define the profit of a solution $\{\tilde{p}_v\}_{v\in V}$ of the generalized problem $\tilde{\Pi} = (\tilde{G}, \tilde{b})$ as

$$\text{profit}_{\tilde{\Pi}}(\tilde{p}) = \sum_{u,v} \sum_{e\in E(u,v)} w_e \begin{cases} \tilde{p}_u + \tilde{p}_v, & \text{if } \tilde{p}_u + \tilde{p}_v \le \tilde{b}_e; \\ 0, & \text{otherwise;} \end{cases}$$

here $E(u, v)$ denotes the set of edges going from u to v. We shall show that the Generalized Graph Pricing problem, even if we allow budgets and weights to be exponential in the number of vertices, is not harder than the standard Graph Pricing problem.

The Generalized Graph Pricing problem is a special case of the general constraint satisfaction problem with constraints depending on two variables (MAX 2GCSP). In our case, the variables are vertices; the constraints or payoff functions are functions

$$f_{\tilde{b}_e}(\tilde{p}_u, \tilde{p}_v) = \begin{cases} \tilde{p}_u + \tilde{p}_v, & \text{if } \tilde{p}_u + \tilde{p}_v \leq \tilde{b}_e; \\ 0, & \text{otherwise.} \end{cases}$$

Strictly speaking, prices can be arbitrary nonnegative real numbers, and thus the domain is infinite. However, if all budgets are positive integers in the range from 1 to B, then the prices in the optimal solution are semi-integral numbers in the range from 0 to B.

Lemma 2.1. *Consider an instance $\Pi = (G, b)$ of the Generalized Graph Pricing problem. Suppose that the budgets $\{b_e\}$ are integers in the range from 1 to B, then prices in one of the optimal solutions are semi-integral numbers in the range from 0 to B.*

Proof sketch. Consider an arbitrary optimal solution $\{p_v\}_{v \in V}$. Let E' be the set of satisfied edges:

$$E' = \cup_{u,v} \{e \in E(u,v) : p_u + p_v \leq b_e\}.$$

Then $\{p_v\}_{v \in V}$ is a solution of the LP: maximize $\sum_{u,v} \sum_{e \in E' \cap E(u,v)} p_u + p_v$ subject to $p_u + p_v \leq b_e$ for all u, v, and $e \in E' \cap E(u,v)$. The LP is semi-integral and thus either all the p_v's are semi-integral numbers or another solution with the same objective value is semi-integral. □

Since we consider only problem instances with integral budgets, we shall assume that all prices are semi-integral. Then the domain size equals $2B + 1$. Note that we could reduce the domain size even further to $O(\log_{(1+\varepsilon)} B) = O(\log(B)/\varepsilon)$ by rounding prices down to powers of $(1 + \varepsilon)$. This reduces the profit of the solution, but by no more than a factor of $(1 + \varepsilon)$.

We now show how to transform an arbitrary Generalized Graph Pricing instance $\tilde{\Pi} = (\tilde{G}, \tilde{b})$ to an unweighted Graph Pricing instance $\Pi = (G, b)$ without parallel edges. We use a relatively standard probabilistic construction that works for arbitrary constraint satisfaction problems. Without loss of generality we assume that the maximum weight is 1.

Input: an instance of Generalized Graph Pricing problem $\tilde{\Pi} = (\tilde{G} = (\tilde{V}, \tilde{E}), \tilde{b})$; a positive ε
Output: an unweighted instance of the Graph Pricing problem $\Pi = (G = (V, E), b)$

- Let m be the total number of edges in the graph $\tilde{G}$; let w be the minimum (non-zero) edge weight.
- Set $N = \lceil m/(w\varepsilon) \rceil^4$.
- For every vertex v of the graph $\tilde{G}$, create N new vertices $v_1, \ldots, v_N$ in the graph G.
- For every edge e between vertices u and v add an unweighted edge between u_i and v_j with probability $\alpha_e = \varepsilon w_e / m$. Set the budget of the new edge to be $b_{(u_i, v_j)} = \tilde{b}_e$. We call this edge a copy of e.
- If an edge (u_i, v_j) is a copy of e and e' $(e \neq e')$ then remove (u_i, v_j) from G.

Lemma 2.2. *Consider an instance of the Generalized Graph Pricing problem* $\tilde{\Pi} = (\tilde{G} = (\tilde{V}, \tilde{E}), \tilde{b})$ *and an instance of the of the Graph Pricing problem* $\Pi = (G = (V, E), b)$ *obtained via the reduction above. Let* $\gamma = \varepsilon N^2/m$*. Then* G *is an unweighted graph without parallel edges; and with probability* $1 - e^{-N}$*,*

$$\frac{OPT_{\Pi}}{\gamma \, OPT_{\tilde{\Pi}}} = 1 + O(\varepsilon).$$

We defer the proof to the Appendix.

Corollary 2.3. *Fix a positive integer* B*. Suppose that it is NP-hard to approximate the Generalized Graph Pricing problem within a factor of* ρ *if all budgets are bounded by* B*. Then for every positive* ε*, it is NP-hard to approximate the Graph Pricing problem within a factor* $(1 - O(\varepsilon))\rho$*.*

Proof. Consider an instance $\tilde{\Pi} = (\tilde{G}, \tilde{b})$ of the Generalized Graph Pricing problem with budgets bounded by B. Let m be the number of edges in the graph $\tilde{G}$. Rescale all weights so that the maximum weight equals 1. Remove all edges with weight less than $\varepsilon m/B$. This decreases OPT_{Π} by at most ε. We now transform the instance $\tilde{\Pi}$ to Π using the reduction from Lemma 2.2. By Lemma 2.2, $OPT_{\tilde{\Pi}} = (1 + O(\varepsilon))OPT_{\Pi}/\gamma$. Thus it is NP-hard to approximate the Graph Pricing problem within a factor $(1 - O(\varepsilon))\rho$. □

Theorem 2.4. *Suppose that it is weakly NP-hard to approximate the Generalized Graph Pricing problem within a factor of* ρ *(i.e. it is NP-hard to approximate the problem within a factor of* ρ *when the budgets and weights can be exponentially large in the problem size). Then, assuming the Unique Games Conjecture, for every positive* ε*, it is NP-hard to approximate the Graph Pricing problem within a factor* $(1 - O(\varepsilon)\rho)$*.*

Proof. We show that there exists a finite set of budgets $\mathcal{B}$ such that if we require all budgets to be from the set $\mathcal{B}$, then the Graph Pricing problem is NP-hard to approximate within a factor of $(1 - O(\varepsilon))\rho$. This is an easy corollary from the recent result of Raghavendra [12]. Raghavendra showed that, assuming the Unique Games Conjecture, the best approximation ratio we can achieve for every 2GCSP problem Λ is at least the integrality gap of the problem Λ (up to any positive constant ε). The problem Λ is defined by a finite set of possible payoff functions and their finite domain.

As mentioned above, we may assume that prices take values in a domain of size $O(\log(B)/\varepsilon)$. Write the standard assignment SDP relaxation for the Generalized Graph Pricing problem (see e.g. Raghavendra [12] SDP (I)). This SDP can be solved in polynomial time. Thus its integrality gap is $(1 - O(\varepsilon))\rho$. Fix an integrality gap example with gap $(1 - O(\varepsilon))\rho$. Let $\mathcal{B} = \{1, \dots, B\}$ be the set containing all budgets from this example. We now consider MAX 2GSP with the set of payoff functions $\{f_b\}_{b \in \mathcal{B}}$ and domain $\{0, 1/2, 1, \dots, B\}$. Its integrality gap is at least $(1 - O(\varepsilon))\rho$. Thus by Raghavendra's theorem [12], it is NP-hard to approximate this MAX 2GCSP problem within a factor $(1 - O(1))\rho$. However, this MAX 2GCSP problem is just the Generalized Graph Pricing problem with budgets bounded by the constant B. □

3 Reduction from Maximum Acyclic Subgraph

We introduce a new problem, which we call Restricted Maximum Acyclic Subgraph. We are given a graph $G = (V, A)$ and a collection of disjoint label sets $S_v \subset \mathbb{N}$ for all vertices v. The goal is to assign a label l_v from the set $S_v \cup \{0\}$ to every vertex v so as to maximize the number of arcs $(u,v) \in A$ for which $l_u < l_v$. The value of a solution is the number of such arcs. We denote the value of the solution $\{l_v\}_{v\in V}$ by $\text{value}_{(G,S)}(l)$; we denote the value of the optimal solution by $OPT_{(G,S)}$.

We now reduce the Restricted Maximum Acyclic Subgraph problem to the Generalized Graph Pricing problem. Given an arbitrary Restricted Maximum Acyclic Subgraph instance $G = (V, A)$, $\{S_v\}_v$ we construct an instance of the Generalized Graph Pricing problem $\Pi = (H, b)$ as follows. The vertices of the graph $H = (V, E)$ are the vertices of the graph G. The edges are triples $(u, v)_l$, where $(u, v) \in A$ and $l \in S_v$. The edge $(u, v)_l$ goes from u to v, has weight M^{-l} and budget $M^l(1 + 1/M)$, where M is a sufficiently large number we specify later. It is convenient to think that the edges are directed; whenever we write $(u, v)_l$ we mean that $(u, v) \in A$. The profit of a solution $\{p_v\}_{v\in V}$ equals

$$\text{profit}_\Pi(p) = \sum_{\substack{(u,v)_l\in E\\ p_u+p_v\le M^l(1+1/M)}} M^{-l}(p_u + p_v).$$

We define the principal profit of the solution as

$$\sum_{\substack{(u,v)_l\in E\\ p_u+p_v\le M^l(1+1/M)}} M^{-l}p_v;$$

and a principal profit of an edge $(u, v)_l$ as $M^{-l}p_v$. We say that a solution $\{p_v\}_{v\in V}$ is *canonical* if

$$p_v \in \{0\} \cup \left\{M^l : l \in S_v\right\}$$

for all v. Every solution $\{l_v\}_{v\in V}$ of the Restricted Maximum Acyclic Subgraph problem corresponds to a canonical solution of the Generalized Graph Pricing problem:

$$p_v = \begin{cases} M^{l_v}, & \text{if } l_v \neq 0;\\ 0, & \text{otherwise.}\end{cases}$$

The principal profit of this solution satisfies

$$\sum_{\substack{(u,v)_l\in E\\ p_u+p_v\le M^l(1+1/M)}} M^{-l}p_v \ge \sum_{\substack{(u,v)\in A\\ p_u+p_v\le M^{l_v}(1+1/M)}} M^{-l_v}\cdot M^{l_v} = \sum_{(u,v)\in A} \begin{cases} 1, & \text{if } l_v > l_u;\\ 0, & \text{otherwise;}\end{cases} = \text{value}_{(G,S)}(l).$$

Thus $\text{profit}_\Pi(p) \ge \text{value}_{(G,S)}(l)$; and $OPT_\Pi \ge OPT_{(G,S)}$. We now show that OPT_Π cannot be much bigger than $OPT_{(G,S)}$. First, we show that the principal profit of every solution almost equals the total profit.

Lemma 3.1. *The profit of an arbitrary solution* $\{p_v\}_{v\in V}$ *of the Generalized Graph Pricing problem* $\Pi = (H, b)$ *defined above is bounded as follows:*

$$\text{profit}_\Pi(p) \equiv \sum_{\substack{(u,v)_l\in E\\ p_u+p_v\le M^l(1+1/M)}} M^{-l}(p_u + p_v) \le \sum_{\substack{(u,v)_l\in E\\ p_u+p_v\le M^l(1+1/M)}} M^{-l}p_v + 2n.$$

Proof. We need to show that

$$\sum_{\substack{(u,v)_l \in E \\ p_u + p_v \le M^l(1+1/M)}} M^{-l} p_u \le 2n.$$

Fix a vertex u. All its outgoing edges have distinct weights and all weights are powers of M. Thus the sequence $M^{-l}p_u$ (where $(u,v)_l \in E$; $p_u \le M^l(1+1/M)$) is a subsequence of a geometric progression with the largest term at most $(1+1/M)$. Hence

$$\sum_{\substack{v:(u,v)_l \in E \\ p_u + p_v \le M^l(1+1/M)}} M^{-l} p_u \le (1+1/M) \sum_{l=0}^{\infty} M^{-l} \le 2.$$

□

We now show how every Generalized Graph Pricing solution can be transformed into a canonical solution.

Lemma 3.2. *For every solution $\{p_v\}_v$ of the problem Π defined above there exists a canonical solution $\{p'_v\}_v$ with the principal profit*

$$\sum_{\substack{(u,v)_l \in E \\ p'_u + p'_v \le M^l(1+1/M)}} M^{-l} p'_v \ge \sum_{\substack{(u,v)_l \in E \\ p_u + p_v \le M^l(1+1/M)}} M^{-l}(p_u + p_v) - (m/M + 2n).$$

Proof. Define

$$p'_v = \begin{cases} M^l, & \text{if } M^{l-1}(1+1/M) < p_v \le M^l(1+1/M) \text{ for some } l \in S_v \\ 0, & \text{otherwise} \end{cases}$$

We compare the principal profit of $\{p'_v\}_{v\in V}$ with the principal profit of $\{p_v\}_{v\in V}$. Consider an edge $(u,v)_l$ with a nonnegative contribution to the profit of $\{p_v\}_{v\in V}$. Then $p_u + p_v \le M^l(1+1/M)$ and both $p'_u, p'_v \le M^l$. Moreover, since $l \in S_v$ and thus $l \notin S_u$ (the sets S_u and S_v are disjoint), $p'_u \le M^{l-1}$. Therefore $p'_u + p'_v \le M^l(1+1/M)$. If $p_v > M^{l-1}(1+1/M)$, then $p'_v = M^l$; and the principal profit of the edge is $M^{-l}p'_v = 1$. If $p_v \le M^{l-1}(1+1/M)$, then $M^{-l}p_v < 2/M$. Hence, the difference $M^{-l}p_v - M^{-l}p'_v$ is always less than $2/M$. We get

$$\sum_{\substack{(u,v)_l \in E \\ p_u + p_v \le M^l(1+1/M)}} M^{-l} p_v \le \sum_{\substack{(u,v)_l \in E \\ p'_u + p'_v \le M^l(1+1/M)}} M^{-l} p'_v + 2m/M;$$

and by Lemma 3.1,

$$\sum_{\substack{(u,v)_l \in E \\ p_u + p_v \le M^l(1+1/M)}} M^{-l}(p_u + p_v) \le \sum_{\substack{(u,v)_l \in E \\ p'_u + p'_v \le M^l(1+1/M)}} M^{-l}(p'_u + p'_v) + m/M + 2n.$$

□

Theorem 3.3. *Consider an instance $(G = (V, A),\, S)$ of the Restricted Maximum Acyclic Subgraph problem. Let $\Pi = (H = (V, E), b)$ be the instance of the Graph Pricing problem obtained through the reduction described above. Then*

$$OPT^{RMAS}_{(G,S)} + 2m/M + 2n \geq OPT_\Pi \geq OPT^{RMAS}_{(G,S)}. \tag{1}$$

Proof. We have already proved that $OPT_\Pi \geq OPT^{RMAS}_{(G,S)}$. Thus we only need to prove the first inequality. Consider an arbitrary solution $\{p_v\}_{v \in V}$. By Lemma 3.2 there exists a canonical solution $\{p'_v\}_v$ with the principal profit

$$\sum_{\substack{(u,v)_l \in E \\ p'_u + p'_v \leq M^l(1+1/M)}} M^{-l} p'_v \geq \sum_{\substack{(u,v)_l \in E \\ p_u + p_v \leq M^l(1+1/M)}} M^{-l}(p_u + p_v) - (m/M + 2n).$$

Set labels l_u as follows: $l_u = \log_M p_u$ if $p_u \neq 0$; and $l_u = 0$ otherwise. If the principal profit of an edge $(u, v)_l$ is greater than $1/M$ then $p_v = M^l$ and $p_u \leq M^{l-1}$. Thus $l_u \leq l_v$ and the arc (u, v) contributes 1 to the value of solution. We get

$$OPT^{RMAS}_{(H,S)} + m/M \geq \sum_{\substack{(u,v)_l \in E \\ p'_u + p'_v \leq M^l(1+1/M)}} M^{-l} p'_v \geq \sum_{\substack{(u,v)_l \in E \\ p_u + p_v \leq M^l(1+1/M)}} M^{-l}(p_u + p_v) - (m/M + 2n).$$

Hence

$$OPT^{RMAS}_{(H,S)} + 2m/M + 2n \geq OPT_G.$$

□

3.1 UG Hardness

Theorem 3.4. *Assuming the Unique Games Conjecture, it is NP-hard to approximate the Graph Pricing problem within a factor $2 - \varepsilon$, for every positive ε.*

Proof. Guruswami, Manokaran, and Raghavendra [8] showed that it is NP-hard to approximate the Maximum Acyclic Subgraph problem within a factor of $2 - \varepsilon$. Observe that the Maximum Acyclic Subgraph problem is a special case the Restricted Maximum Acyclic Subgraph problem, where sets S_v are chosen so that any ordering of vertices is possible. Hence, by Theorem 3.3 we can transform any graph G to an instance of the Generalized Graph Pricing problem Π satisfying

$$OPT^{MAS}_G + 2m/M + 2n \geq OPT_\Pi \geq OPT^{MAS}_G,$$

where OPT^{MAS}_G denotes the size of maximum acyclic subgraph in G. Assume for a moment that $2m/M + 2n \leq \varepsilon\, OPT^{MAS}_G$. Then $(1+\varepsilon)OPT^{MAS}_G \geq OPT_\Pi \geq OPT^{MAS}_G$; and thus the Generalized Graph Pricing problem is NP-hard to approximate within a factor of $2 - O(\varepsilon)$. Theorem 2.4 implies that the Graph Pricing problem is then also NP-hard to approximate within a factor of $2 - O(\varepsilon)$.

We now take care of the term $2m/M + 2n$. We replace every vertex v in G by $K = \lceil 1/\varepsilon \rceil$ new vertices $v_1, \dots, v_K$ and every edge (u, v) with K^2 edges (u_i, v_j). The number of vertices in the graph increases K times; the number of edges and the size of the maximum acyclic subgraph increases exactly K^2 times. (Since vertices $v_1, \dots v_K$ have exactly the same neighbors they can be arranged consecutively in the optimal solution.) Pick $M = mK^2$. Then $2m/M + 2n \leq 2 + \varepsilon OPT^{MAS}_G$. □

3.2 LP Integrality Gap

We study the following LP relaxation:

$$\sum_{(u,v)\in E} \sum_{\substack{q,s \\ q+s\leq b_{(u,v)}}} (q+s) y_{uv}(q,s),$$

subject to

$$\sum_q x_u(q) = 1 \qquad \text{for all } u \tag{2}$$

$$\sum_s y_{uv}(q,s) = x_u(q) \qquad \text{for all } u, v, q \tag{3}$$

$$y_{uv}(q,s) = y_{vu}(q,s) \qquad \text{for all } u, v, q, s \tag{4}$$

$$0 \leq x_u(q) \leq 1 \qquad \text{for all } u, q \tag{5}$$

$$0 \leq y_{uv}(q,s) \leq 1 \qquad \text{for all } u, v, q, s \tag{6}$$

In the *intended* integral solution, each $x_u(q)$ is the indicator variable of the event "the vertex u has price q," i.e., $x_u(q) = 1$, if $p_u = q$; $y_{uv}(q,s) = 1$, if u has price s, v has price q; and is equal to 0, otherwise. It is easy to see that in the intended integral solution all the constraints are satisfied. As before we assume that budgets $\{b_e\}_{e\in V}$ are integral and indexes q,s take semi-integral values in the range 0 to $\max_e b_e$. Note that the LP upper bound on the optimal solution is stronger than the combinatorial upper bound of Balcan and Blum (see the introduction). Indeed, for all u, we have

$$\sum_v \sum_{q,s:q+s\leq b_{(u,v)}} y_{uv}(q,s) \times q \leq \sum_q x_u(q) \sum_{v:q\leq b_{(u,v)}} q \leq \max_q \sum_v f_{(u,v)}(q,0) = R(u).$$

Our LP integrality gap example is based on the construction of Alon, Bollobàs, Gyàrfàs, Lehel, and Scott [1].

Theorem 3.5 (Alon et al. [1]). *There exists a directed acyclic graph G having m edges and $n = o(m)$ vertices, such that every directed cut of G contains at most $(1/4 + o(1))m$ edges.*

Theorem 3.6. *The integrality gap of the LP is $(4 - \varepsilon)$, for every positive ε.*

Proof. Let $G = (V, A)$ be the graph of Alon et al. [1]. We order the vertices of G in the reverse topological order. For every $v \in V$, let $o_v \in \{1, \ldots, n\}$ be the position of the vertex v in the ordering. Then if $(u, v) \in A$, $o_u > o_v$. Fix an integer parameter T. Construct an instance of the Restricted Maximum Acyclic Subgraph problem on graph G. Set

$$S_v = \{o_v \times T, o_v \times T + 1, \ldots, o_v \times T + T - 1\}.$$

For every edge $(u, v) \in A$, valid assignments of labels l_u and l_v that satisfy the inequality $l_u < l_v$ are $l_u = 0$; $l_v \in S_v$. Thus the value of any solution $\{l_v\}_{v\in V}$ equals the size of the directed cut between the sets $\{u : l_u = 0\}$ and $\{v : l_v \in S_v\}$. Therefore, the optimal value of the solution is at most $(1/4 + o(1))m$. We transform (G, S) to an instance of the Generalized Graph Pricing problem (using Theorem 3.3) and then to an unweighted instance of the Graph Pricing problem (using

Lemma 2.2). The profit of the optimal solution of the obtained problem $\Pi = (H = (V_H, E_H), b)$ is at most $(1/4 + O(\varepsilon))m \times \gamma N^2$ (if we choose M to be sufficiently large).

We now describe an LP solution of value $(1 - 1/T)m \times \gamma N^2$. Recall that the vertices of H are pairs in $V \times \{1, \ldots, N\}$ denoted v_i. The set of edges is a random subset of triples $(u_i, v_j)_l$, where $(u, v) \in A$, $l \in S_v$. The budget of $(u_i, v_j)_l$ is M^l. The probability that the edge $(u_i, v_j)_l$ is present in the graph is $\alpha'_{(u_i,v_j)_l} = \gamma/(M^l N^2)(1 - O(\varepsilon))$. We choose edges, so that the graph does not have parallel edges.

Set LP variables $x_{v_i}(M^l) = 1/T$, and $x_v(0) = 1/T$ for all vertices v_i and $l \in S_v$. Note that S_v contains exactly $T - 1$ elements, thus $x_{v_i}(0) + \sum_l x_{v_i}(M^l) = 1$. For every edge $(u, v)_l$ set $y_{u_i v_j}(0, M^l) = 1/T$. Set all other $y_{u_i v_j}(s, q)$ arbitrary to satisfy the LP constraints (e.g. $y_{uv}(M^l, 0) = 1/T$; $y_{uv}(M^{l'}, M^{l'}) = 1/T$ for $l' \neq l$).

If an edge $(u_i, v_j)_l$ is present in the graph, then its contribution to the LP objective function is at least $(0 + M^l) \times y_{u_i v_j}(0, M^l) = M^l/T$. Thus for every $(u, v) \in A$, the expected contribution of all edges $(u_i, v_j)_l$ is at least

$$\sum_{l \in S_v} \sum_{1 \leq i,j \leq N} \frac{(1 - O(\varepsilon))\gamma}{M^l N^2} \frac{M^l}{T} = (T - 1) \times N^2 \times \frac{(1 - O(\varepsilon)) \times \gamma}{M^l N^2} \frac{M^l}{T} = (1 + O(\varepsilon)) \times \frac{T - 1}{T} \times \gamma.$$

We have proved that for every positive ε, there exists a graph with the cost of the optimal solution at most $\gamma m/4 \times (1 + O(\varepsilon))$ and the cost of the LP at least $\gamma m \times (1 - O(\varepsilon))$. Hence the integrality gap is $4 - O(\varepsilon)$. □

Remark 3.7. *A similar construction shows that the problem is (unconditionally) at least as hard as MAX CUT, which as was shown by Håstad [10] cannot be approximated better than within a factor of* 17/16 *(unless $P = NP$).*

References

[1] N. Alon, B. Bollobàs, Gyàrfàs, A. J. Lehel, and A. Scott. Maximum directed cuts in acyclic digraphs. *J. Graph Theory*, 55:1–13, 2007.

[2] M.-F. Balcan and A. Blum. Approximation algorithms and online mechanisms for item pricing. *Theory of Computing*, 3:179–195, 2007.

[3] P. Briest and P. Krysta. Single-minded unlimited supply pricing on sparse instances. In *Proceedings, ACM-SIAM Symposium on Discrete Algorithms*, pages 1093–1102, 2006.

[4] E. D. Demaine, U. Feige, M. Hajiaghayi, and M. R. Salavatipour. Combination can be hard: Approximability of the unique coverage problem. In *Proceedings, ACM-SIAM Symposium on Discrete Algorithms*, pages 162–171, 2006.

[5] K. Elbassioni, R. Sitters, and Y. Zhang. A quasi-ptas for profit-maximizing pricing on line graphs. In *Proceedings, European Symposium on Algorithms*, pages 451–462, 2007.

[6] A. Goldberg, J. Hartline, A. Karlin, M. Saks, and A. Wright. Competitive auctions. *Games and Economic Behavior*, 55(2):242–269, 2006.

[7] V. Guruswami, J. Hartline, A. Karlin, D. Kempe, C. Kenyon, and F. McSherry. On profit-maximizing envy-free pricing. In *Proceedings, ACM-SIAM Symposium on Discrete Algorithms*, pages 1164–1173, 2005.

[8] V. Guruswami, R. Manokaran, and P. Raghavendra. Beating the random ordering is hard: Inapproximability of maximum acyclic subgraph. In *focs*, pages 573–582, 2008.

[9] J. Hartline and V. Koltun. Near-optimal pricing in near-linear time. In *Proceedings, Workshop on Algorithms and Data Structures*, pages 422–431, 2005.

[10] J. Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.

[11] R. Krauthgamer, A. Mehta, and A. Rudra. Pricing commodities, or how to sell when buyers have restricted valuations. In *Proceedings, Workshop on Approximation and Online Algorithms*, 2007.

[12] P. Raghavendra. Optimal algorithms and inapproximability results for every csp? In *STOC*, pages 245–254, 2008.

A Proof of Lemma 2.2

Proof of Lemma 2.2. Consider an edge e between two vertices u and v in $\tilde{G}$. We add a copy of e between u_i and v_j at step 4 with probability α_e. The probability that we remove the edge at the last step is less than $\alpha_e \times m\alpha_e \le \varepsilon$. Thus the probability β_e that the obtained graph G has the edge (u_i, v_j) is between $(1-\varepsilon)\alpha_e$ and α_e.

Let $\mathcal{V}^u$ and $\mathcal{V}^v$ be arbitrary subsets of $\{u_i : 1 \le i \le N\}$ and $\{v_j : 1 \le j \le N\}$ respectively. Denote by $E_e(\mathcal{V}^u, \mathcal{V}^v)$ the set of copies of e going from $\mathcal{V}^u$ to $\mathcal{V}^v$. The expected size of $E_e(\mathcal{V}^u, \mathcal{V}^v)$ is $\beta_e |\mathcal{V}^u|\,|\mathcal{V}^v|$. By a Bernstein or Chernoff type inequality,

$$\Pr\left(\big| |E_e(\mathcal{V}^u, \mathcal{V}^v)| - \beta_e |\mathcal{V}^u| \cdot |\mathcal{V}^v| \big| \le 4N^{3/2} \right) \le 2e^{-\frac{16N^3}{2(\beta_e |\mathcal{V}^u| \cdot |\mathcal{V}^v| + N^{3/2}/3)}} \le e^{-4N}.$$

The number of ways we can choose sets $\mathcal{V}^u$ and $\mathcal{V}^v$ is 2^{2N}. Thus, by the union bound, with probability at least $1 - e^{-2N}$, for all $\mathcal{V}^u \subset \{u_i : 1 \le i \le N\}$ and $\mathcal{V}^v \subset \{v_j : 1 \le j \le N\}$,

$$\big| |E_e(\mathcal{V}^u, \mathcal{V}^v)| - \beta_e |\mathcal{V}^u| \cdot |\mathcal{V}^v| \big| \le 4N^{3/2}.$$

Moreover, since the number of edges m is less than e^N, with probability $1 - e^{-N} > 0$, for all u, v, $e \in E(\mathcal{V}^u, \mathcal{V}^v)$, $\mathcal{V}^u \subset \{u_i : 1 \le i \le N\}$ and $\mathcal{V}^v \subset \{v_j : 1 \le j \le N\}$,

$$\big| |E_e(\mathcal{V}^u, \mathcal{V}^v)| - \beta_e |\mathcal{V}^u| |\mathcal{V}^v| \big| \ge 4N^{3/2}. \tag{7}$$

We fix one of the random instances satisfying this condition. Given an arbitrary semi-integral solution p_{v_i} of the problem Π, we define a probabilistic solution of the original problem $\tilde{\Pi}$ as follows: for every vertex v pick a random i from 1 to N and set $\tilde{p}_v = p_{v_i}$. For all v and all semi-integral q, let $\mathcal{V}^v_q = \{v_i : \tilde{p}_{v_i} = q\}$. The probability that we assign price q to u and s to v equals $|\mathcal{V}^u_q \times \mathcal{V}^v_s|/N^2$. Thus the expected profit of $\tilde{p}$ equals

$$\mathbb{E}\left[\text{profit}_{\tilde{\Pi}}(\tilde{p})\right] = \sum_{u,v} \sum_{e \in E(u,v)} \sum_{q,s} \frac{|\mathcal{V}^u_q||\mathcal{V}^v_s|}{N^2} \times w_e f_{b_e}(q,s).$$

The profit of p equals

$$\mathrm{profit}_{\Pi}(p) = \sum_{u,v} \sum_{e \in E(u,v)} \sum_{q,s} |E_e(\mathcal{V}_q^u, \mathcal{V}_s^v)| \times f_{b_e}(q,s).$$

Thus,

$$\mathrm{profit}_{\Pi}(p) - \gamma \cdot \mathbb{E}\left[\mathrm{profit}_{\tilde{\Pi}}(\tilde{p})\right] = \sum_{u,v} \sum_{e \in E(u,v)} \sum_{q,s} (|E_e(\mathcal{V}_q^u, \mathcal{V}_s^v)| - \alpha_e |\mathcal{V}_q^u||\mathcal{V}_s^v|) \times f_{b_e}(q,s).$$

By (7),

$$\begin{aligned}\mathrm{profit}_{\Pi}(p) - \gamma \cdot \mathbb{E}\left[\mathrm{profit}_{\tilde{\Pi}}(\tilde{p})\right] &\le \textstyle\sum_{u,v} \sum_{e \in E(u,v)} \sum_{q,s} (|E_e(\mathcal{V}_q^u, \mathcal{V}_s^v)| - \beta_e |\mathcal{V}_q^u||\mathcal{V}_s^v|) \times f_{b_e}(q,s) \\ &\le m \times 4N^{3/2} \times \max_e b_e \le \varepsilon \times \gamma w \max_e b_e.\end{aligned}$$

Since $OPT_{\Pi} \ge w \max_e b_e$,

$$OPT_{\Pi} \ge \gamma \; OPT_{\tilde{\Pi}} (1 + O(\varepsilon)).$$

Similarly, given a solution $\{p_v\}_{v \in V}$ of the problem $\tilde{\Pi}$, we define a solution of Π as $\tilde{p}_{v_i} = p_v$. Then

$$\begin{aligned}(1-\varepsilon)\frac{\varepsilon N^2}{m} \cdot \mathrm{profit}_{\tilde{\Pi}}(\tilde{p}) - \mathrm{profit}_{\Pi}(p) &\le \textstyle\sum_{u,v} \sum_{e \in E(u,v)} \sum_{q,s} \left||E_e(\mathcal{V}_q^u, \mathcal{V}_s^v)| - \beta_e |\mathcal{V}_q^u||\mathcal{V}_s^v|\right| \times f_{b_e}(q,s) \\ &\le m \times 4N^{3/2} \times \max_e b_e \le \varepsilon \times \gamma w \max_e b_e.\end{aligned}$$

Thus,

$$(1-\varepsilon) OPT_{\tilde{\Pi}} \ge \gamma \; OPT_{\Pi} (1 + O(\varepsilon)).$$

□

B Hardness of the bipartite case

Balcan and Blum achieve a 4-approximation by reducing the general problem to the bipartite case, and they note that any improvement over the trivial 2-approximation for the bipartite case would immediately improve the 4-approximation for the general case. Here we show that the bipartite case is APX-hard, which to the best of our knowledge was previously unknown.

Theorem B.1. *The Graph Pricing Problem in the bipartite case is APX-hard.*

Proof. We reduce the well-known APX-hard problem MAX CUT to graph pricing in a bipartite graph. Let $G = (V, E)$ be a (general) graph. For each vertex u in V we will construct a vertex u in our bipartite graph $G' = (V_1, V_2, E')$. For convenience we will refer to these corresponding nodes using the same names. All the original nodes in G will be on the same side of G', say V_1. For each edge $(u, v) \in E$ we construct the gadget shown in Figure 1.

We now show that if u and v both charge 0 or both charge 1, the most profit that can be gained from the gadget is 8, whereas profit 9 can be obtained if one charges 0 and the other charges 1. We will later show how to make sure each of u and v charges either 1 or (almost) 0.

First suppose that both u and v charge 0. Note that the total budget of edges in the gadget is 10. Thus if any edge with budget 2 has its budget exceeded, the profit is at most 8 and we are done. On the contrary, some edge's budget of 1 must be exceeded to collect more than 8, as follows. Suppose w and x together charge at most 1 and also y and z together charge at most 1. Then these 4 nodes collect a total of at most 4 out of a possible 6 from the 4 edges on the cycle, since each collects twice its charge from those edges, and the total profit is again at most 8. So now assume some budget of 1 is exceeded. By symmetry we may assume the budget of edge (w, x) is exceeded. If edge (y, z) also has its budget exceeded, again the remaining budget is only 8 and we are done. Thus z can charge at most 1, so edge (z, v) pays at most 1 from its budget of 2, and in this last case, again the total profit is at most 8.

Now suppose u and v each charge 1. Again note that if any edge with budget 2 has its budget exceeded, the remaining budget is 8 and we are done, that some budget of 1 must be exceeded to collect more than 8, and by symmetry we may assume it is (w, x). w and z charge some $p_w \leq 1$ and $p_z \leq 1$, y charges some $p_y \leq 1 - p_z$, and x charges some $p_x \leq 2 - p_z$. We may assume $p_w = 1$, $p_y = 1 - p_z$, and $p_x = 2 - p_z$ since these can only increase profit and will not cause any (more) budgets to be exceeded. Thus the total profit, vertex by vertex in alphabetical order, is $1 + 1 + 2 + (2 - p_z) + 2(1 - p_z) + 3p_z = 8$.

Finally, suppose u charges 0 and v charges 1. The total profit can be 9 as follows and by symmetry the same will hold if these are reversed. It is easy to check that the total profit is 9 if w charges 2, x and z charge 1, and y charges 0.

Unfortunately, profit greater than 8 can be extracted from our gadget via fractional charges at u and/or v. To ensure that each node corresponding to a node in the original graph charges either 1 or ε for some small ε, we use the gadget displayed in Figure 2. We will show that for any solution, we can find a solution that satisfies this condition and whose value is at least that of the given solution.

First we show that we can assume node a charges 0, as follows. We claim nodes a, b, c, and d can together collect at most 4. It is easy to see that this can be achieved by those nodes charging, respectively, 0, 1, 1, and 0. Suppose to the contrary that a charges some $p_a \leq 1$, since it is trivial that a charge greater than 1 yields the claim. Note that if any edge's budget exceeded, the remaining budgets are 4 or less and we are done. Then, in order, the nodes collect at most $2p_a + 2(1 - p_a) + p_c + (1 - p_c) = 4$, since a charge by b greater than $1 - p_a$ breaks the budget of edge (a, b), and of course c and d must not charge more than a sum of 1. Thus we can assume a, and each of the corresponding nodes in the similar chains with budgets ε, etc., charge 0.

Now suppose v charges a value p_v other than 1 or ε. If $p_v > 1$, then clearly v collects at most 2 and we can reduce p_v to 1, for which v collects 2. If $\varepsilon \leq p_v \leq 1/2$, v collects at most 1, and by changing p_v to ε, v collects 1. Since we only decrease p_v in these cases, we cannot break the budget of (z, v). Now suppose $p_v < \varepsilon$. We increase it to ε to bring the profit up to 1. This may force us to decrease p_z by ε. We will account for this small error later. Finally, suppose $1/2 < p_v < 1$. Suppose we increase p_v by $\delta = 1 - p_v$. We must consider two cases. If $p_v + p_z < 2$, then p_z can remain unchanged and we only increase the total value. If $p_z + p_v = 2$, then we must decrease p_z by δ to avoid breaking the budget of (z, v). But in this case, $p_z > 1$, so z's incident edge with budget 1 is not paying. Thus z loses 2δ and v gains this same amount. (We may need to combine these two cases in case $p_z + \delta > 2$; the details are straightforward.)

Finally note that whether v charges 1 or ε, the gadget contributes a fixed amount of 8 to the total value of the solution. We attach this gadget to both ends of each edge (u, v), and thus if u and v each charge ε or each charge 1, the total value contributed by the combination of gadgets is approximately 24, whereas the value is approximately 25 if one charges 1 and the other ε.

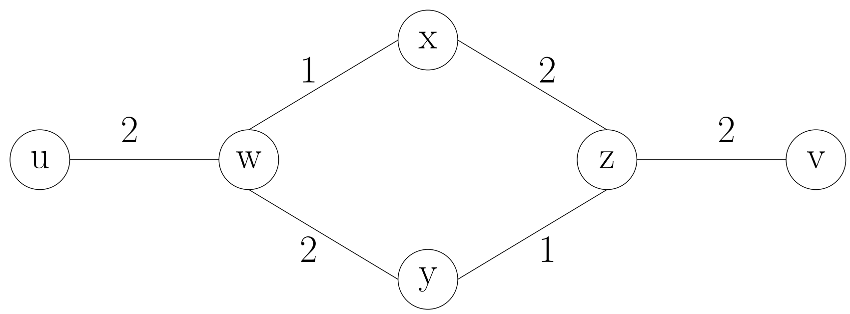

Figure 1: Bipartite pricing gadget. Note that despite the layout, u and v belong to the same side of the bipartition.

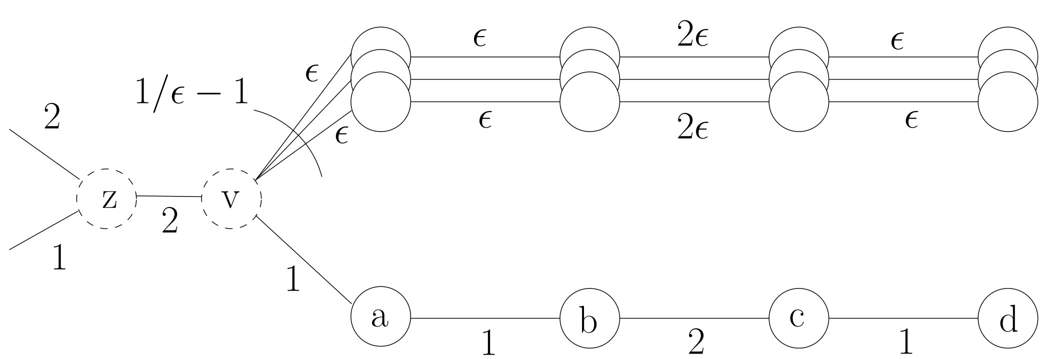

Figure 2: Price-enforcing gadget: node v charges 1 or ε.

For a node u of degree d, both gadgets are replicated for each of its edges. It is easy to see that this still forces each node to charge either 1 or ε.

It should be clear now that (neglecting at most 2ε per edge), from a solution of value $24m + C$ to our bipartite graph pricing instance, we can recover a cut of size C to the original MAX CUT instance, where m is the number of edges. We can make the error insignificant by appropriate choice of ε. By the well-known APX-hardness of MAX CUT, we obtain the APX-hardness of bipartite graph pricing.

□